FORECASTING THE FUTURE WATER HEIGHT OF A RESERVOIR WITH TEMPORAL DATA MINING

Mohamed Saudulla Aabid Rushdi

168265J

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

June 2018

DECLARATION

I declare that this is my own work and this thesis does not incorporate without acknowledgement, any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works.

....

Mohamed Saudulla Aabid Rushdi Date

The above candidate has carried out research for the Masters thesis under my supervision.

....

Dr. Amal Shehan Perera Date

ABSTRACT

Throughout last few decades, flooding has been one of the most expensive natural disasters, which has had an increased impact on human causalities, property damage and rehabilitation. Thus, there is a crucial need in taking necessary actions to avoid or minimize the impact caused by floods to human lives as well as to the stability of the economy of the country. As to the most recent devastating flood experience faced in Sri Lanka in 2016, one of the root causes identified was the unpredictable decision making strategy, which was used to control and manage excess water in reservoirs due to heavy rain. Moreover, reservoirs are a key water storage source in water management, where they are utilized by various sectors for different purposes. Therefore, there is an essential need in taking the best decision in releasing water from reservoirs to minimize the damage caused by the floods and to optimize the utilization of water as a scarce resource.

Many researches have been carried out on the field of data science based on collected data from rivers, reservoirs & tanks to support decision making in water management. Those are mainly focused on classifying water release rules and ranges of a reservoir or tank. Research on forecasting the future water height of reservoirs, with both rainfall and uncertain water inflow due to human intervention, are very limited.

Hence, this research focuses on predicting the future water height of a reservoir, when the water inflow is uncertain and the reservoir receives a significant amount of rainfall. This would allow to minimize the risk of deadly floods by opening the sluice gates in time and to manage the water in an optimum manner for irrigation purposes. This research proposes the most effective set of features to forecast the water height of a reservoir on next three days. Furthermore, it presents the comparison of the performance of different regression models and the effectiveness of applying clustering techniques on top of the regression models. The result obtained from this research demonstrates that, K-Medoids clustering with feed forward artificial neural network model has the best performance in forecasting the reservoir water height when there is a significant amount of rainfall and the water inflow is uncertain.

ACKNOWLEDGEMENT

There is a tremendous amount of hard work and commitment behind the success of this research and I take this opportunity to convey my heartiest gratitude towards all the individuals who were a part of it. It is indeed my privilege to thank each one of them who gave me their valuable support, constant and necessary guidance while giving me the right level of motivation towards reaching this research's ultimate goal.

Firstly, I would like to express my heartfelt gratitude Dr. Amal Shehan Perera, Department of Computer Science and Engineering; as my project supervisor for his remarkable support, supervision and the inspiring advices given throughout, to make this research a success.

While my sincere thanks also go to Dr. Malaka Walpola, Dr. Charith Chitraranjan, Dr. Surangika Ranathunga and Dr. Indika Perera for their valuable feedback, encouragement and advices through time to time evaluations and providing me with the necessary advice and guidance in initiating and carrying out this research.

I also take this opportunity to thank Mr. Sanjaya Rathnayaka, and Mahaweli Authority for providing me the required data and helping me with the domain knowledge. I also should be grateful to darksky.net and timeanddate.com websites, from where I have collected some of the necessary data for the simulation.

Finally, I would like thank to each individual; especially my parents for their love & motivation and my loving wife for her support towards making this research a success during this entire period of time. I would also like to express my thanks to all my colleagues at MSc 16 batch and at Synopsys Lanka (PVT) Ltd.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

Abbreviation	Description
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
AR	Auto-Regressive
ARMA	Auto-Regressive Moving Average
CHRS	Center for Hydrometeorology and Remote Sensing
Day 0	The day from which the future water height is forecasted
Day 1	Following day to a given date (day 0), on which the water height is forecasted
Day 2	Second consecutive day from day 0, on which the water height is forecasted
Day 3	Third consecutive date from day 0, on which the water height is forecasted
GEP	Gene Expression Programming
GIS	Geographic Information System
IWMI	International Water Management Institute
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NN	Neural Network
PCA	Principal Component Analysis
RF-SVR	Reduced Feature - Support Vector Regression
RMSE	Root Mean Square Error
SPI	Standardized precipitation index
SVM	Support Vector Machine
SVR	Support Vector Regression
TDANN	Time Delay Artificial Neural Network